

Preface

Management of enterprise-wide databases and access to managed data is a growing challenge for data processing customers. Databases that have grown over time within various parts of an organization may not share a single view of the data, may have different tools, and may lack consistency in format and meaning. In addition, these large databases may lack a record of changes, may have conflicting rules for data sharing and general data administration, and may present a complex web of management issues such as timeliness, support, integrity, and maintenance.

This issue presents a current view of the technologies and environments for management of large enterprise databases. Seven papers show current experiences and solutions, extensions to the relational model and systems, support for object management systems, distributed databases, and parallel processing for database systems. The experience, knowledge, and technology described is, in turn, the focus of the ongoing development of IBM's Information Warehouse* framework. This framework supports relational and object approaches to data and encompasses an architecture and conforming implementations. We are indebted to N. E. Owens of the IBM Santa Teresa Laboratory, IBM Software Solutions Division, in San Jose, California, for her coordination and management of the preparation of these papers.

Some enterprises are addressing the current complexities of enterprise-wide databases and data by applying innovative approaches to their existing systems and technology. In the first of two papers addressing current solutions, Bustamente and Sorenson describe the management of operational data gathered from order processing and customer mailing at Lands' End, Inc. The resulting corporate database is accessed through a decision support system built using DATABASE 2* (DB2*), in a way that can be considered a precursor to and subset of the Information Warehouse

framework. The paper discusses corporate requirements, database design process and modeling, tool selection, and insights gained during implementation.

Schlatter et al. offer a second paper on current experiences that also treats technical advances intended to pave the way for enterprise-wide database solutions. Their Business Object Management System (BOMS), which is operational today, is intended to support billions of objects of all kinds at one time. It is built on DB2 and Customer Information Control System/Enterprise Systems Architecture (CICS/ESA*), extending the relational database model in ways that avoid the problem of predicting future uses of data and relationships among data. It also provides transparent access to data, portending a future of cooperative database management in which data navigation is no longer an end-user concern.

Turning our attention to technology for the Information Warehouse framework, both a bridge for existing database management systems and support for newer approaches are needed, so that the current and foreseen challenges for databases can be met. The first paper of two on technology presents extensions to relational database management systems (RDBMSs) and models, supporting new capabilities such as extensible types and functions, a global semantic rules subsystem, and performance enhancements. These extensions, as described by Cheng et al., are currently implemented and tested as prototypes for the Structured Query Language (SQL).

In the second paper of this set, Alfred provides a view of the technology for ObjectStore**, an object database management system (ODBMS) that is now part of IBM's object database solution and the database component of IBM's FlowMark* work flow management system. He discusses two paradigm shifts and the leverage they provide: the object-oriented development model and the di-

rect-reference storage model. These models underlie some current database systems, including ObjectStore, and are part of the adoption of increasingly sophisticated object-based software development techniques.

A further set of two papers centers on the design and development of distributed database applications within the context of the Information Warehouse framework. The first, by Singleton and Schwartz, explores direct data access in a heterogeneous systems environment, with the goal of supporting efficient database and systems administration for a large enterprise. Specific software development techniques are presented, along with technical trade-offs that need to be considered in their use.

Leymann and Altenhuber provide a second paper on distributed databases that focuses on the modeling, management, storage, and execution of enterprise-wide business process descriptions, and their importance as information assets. The meta-model supporting this work utilizes weighted, colored, directed graphs of activities, executed by graph navigation. This meta-model is implemented in FlowMark and used for the work flow management component of business processes. Comparisons are made with other methods, such as Petri nets and process models.

The pressure for efficiency and rapid responses to queries in large database systems has made it beneficial to consider parallelism for query processing. With a proposal for using parallelism comes the need to consider secondary characteristics, such as load balancing, pipelining, database design, and parallelization of the query itself. The authors, Mohan et al., show how these questions were resolved and implemented in the chosen relational database management system: DB2 Version 3.

The next issue of the Journal will be a special issue on projects sponsored by IBM Canada's Centre for Advanced Studies.

Gene F. Hoffnagle
Editor

*Trademark or registered trademark of International Business Machines Corporation.

**Trademark or registered trademark of Object Design, Inc.